

# A Method for the Investigation of those Transformations under which the Visual Recognition of a Given Object is Invariant

## I. The Theory

DAVID H. FOSTER

Imperial College of Science and Technology, Applied Optics Section, Department of Physics, London, Great Britain

Received: June 7, 1972

### Abstract

The theory is developed of a method for the investigation of those transformations which have the property that for a given object their action leaves visual recognition of the object invariant.

The technique rests on a visual apparent movement effect known as phi-motion. Two assertions are arrived at. The first is that the existence of phi-motion between an object and transform implies the capacity of the visual system to effect a certain 1-parameter family of local transformations, and, in particular, implies the invariant recognition of the given object under the given transformation. The second is that it is possible to determine whether the observed 1-parameter family is part of a local 1-parameter group. A systematic investigatory program is built up on the basis of these two assertions. Application of this program is claimed to yield data upon the construction of the set of transformations "carried" by the visual system, and the nature of those paths in this set which are preferred by the system for execution.

### Introduction

There is a well known problem associated with Visual Perception: How is it that one can recognize an object, visually, even when it has undergone transformations like rotations, translations, dilations, etc.? Pitts and McCulloch in their classic paper "How We Know Universals" (Pitts and McCulloch, 1947) discuss this phenomenon, and put forward two possible neural schemes which could effect such *invariant recognition* of an object. A review of this study is given in Arbib (1971).

The concern of the work to be reported below is not so much with the general and very difficult problem of constructing neural mechanisms for invariant recognition, as with the initial problem of establishing precisely *which* transformations have this invariance property. A priori, we can certainly make the (almost trivial) assertion that recognition of an object is not invariant under all transformations (even under those which serve to keep the image of the object on the retina). Consider

i) the group of dilations acting (locally) on the retina. One can demagnify a given retinal image to

an arbitrary extent, and, in particular, take it below the limit of visual resolution.

ii) the group of translations acting (locally) on the retina. One can move two visually distinguishable images from one retinal location with a certain associated spatial frequency response function to another retinal location with a much poorer spatial frequency response function, and thereby render the two visually indistinguishable.

In view of these comments, it might appear appropriate that we commence the present study by introducing Zeeman's *tolerance relation* upon the retina (Zeeman, 1962). A formal handling of visual indistinguishability in the present context, however, leads to some complication in presentation, and therefore in order to proceed as quickly as possible, we forego any such attempt at a complete treatment.

In what follows, we put forward the theory of a simple investigatory procedure, by which, it is claimed, one can pick out certain of the transformations under which the recognition of a given object is invariant. Further, in connection with the general problem of building neural mechanisms, it is suggested that the proposed method can provide data upon the composition of the set of transformations "carried" by the system, and, in particular, can reveal when a 1-parameter family of transformations executed by the system is, locally, a 1-parameter group.

The paper immediately after this (Foster, 1972b) describes, in detail, an experiment which applies this investigatory procedure to the special case of the group of rotations  $SO(2)$  acting on a Landolt ring.

### Theory

We first introduce some notation and some definitions.

Let us consider the retina  $V$  as a hemisphere in  $R^3$  ( $R$  is the set of real numbers) and let  $U$  be a (non-

empty) open subset of  $V$  (for example,  $U$  might be the fovea). We shall assume  $U$  to be fixed unless otherwise indicated. A visual stimulus or *object*  $A$  on  $V$  is defined as a mapping (not necessarily continuous) of  $V$  into  $\mathbf{R}$  such that for  $p \in V$ ,  $A(p)$  ( $A(p) \geq 0$ ) gives the illumination on the retina at the point  $p$ . The set of all objects on  $V$  (resp.  $U$ ) will be denoted by  $F(V)$  (resp.  $F(U)$ ). For a given object  $A$ ,  $N(A)$  will stand for the set of points in  $V$  for which  $A(p) \neq 0$ . [We will adopt the convention of writing " $p \in A$ " to indicate that the point  $p$  is to be considered in conjunction with its illumination value  $A(p)$ .]

The group of bijective mappings of  $V$  onto  $V$  will be denoted by  $\text{Bij}(V)$ , and the groupoid of injective mappings of  $U$  into  $V$  by  $\text{Inj}(U, V)$ . [ $\text{Inj}(U, V)$  is a groupoid in that where the compositions of mappings are defined, they follow the group laws.] Given a mapping  $\tau \in \text{Inj}(U, V)$ , the simplest (non-trivial) way of defining its action on elements of  $F(U)$  is to put  $(\tau(A))(q) = A(\tau^{-1}(q))$  for  $A \in F(U)$  and  $q \in \tau(U)$  (1)

which assigns to each point  $q$  of the transformed object the illumination at its preimage.

We have not stipulated that the mappings  $\tau \in \text{Inj}(U, V)$  need be differentiable or even continuous. The introduction of such a restriction, however, does not lead to a very great loss in generality, since, for most mappings of interest (as far as vision is concerned), it should always be possible to find some subset of  $U$  upon which  $\tau$  is suitably smooth. Indeed, one might go further, and assert that for any mapping  $\tau: U \rightarrow V$  it is always possible to find an arbitrarily differentiable mapping  $\tau': U \rightarrow V$  visually indistinguishable from  $\tau$  on all of  $U$ . Henceforth, therefore, only those  $\tau \in \text{Inj}(U, V)$  for which  $\tau: U \rightarrow V$  is a ( $C^\infty$ ) diffeomorphism will be considered.  $\text{Inj}(U, V)$  (and  $\text{Bij}(V)$ ) will from now on be taken to consist solely of such *transformations*.

Returning to the question of defining the action of  $\tau$  on members of  $F(U)$ , we might now consider putting

$$(\tau(A))(q) = \frac{A(p)}{|\det(D\tau(p))|} \quad (2)$$

for  $A \in F(U)$  and  $q = \tau(p) \in \tau(U)$ .

$D\tau(p)$ , which is non-singular, is the derivative of  $\tau$  at the point  $p$ . The RHS of (2) makes sense if, considering  $V$  as part of a sphere  $S^2$  in  $\mathbf{R}^3$ , we introduce local coordinate systems  $(U_q, \varphi_q)$  and  $(U_p, \varphi_p)$  at  $q$  and  $p$  respectively ( $U_q$  is an open neighbourhood of  $q$ , and  $\varphi_q: U_q \rightarrow \varphi_q(U_q) \subset \mathbf{R}^2$  is a homeomorphism; similarly for  $p$ ) such that  $U_q = \varrho(U_p)$  for some  $\varrho \in O(3)$ , and  $\varphi_q(q') = \varphi_p(\varrho^{-1}(q'))$  for all  $q' \in U_q$ . [ $O(3)$  is the orthogonal group acting on  $S^2$ .]

The motivation behind definition (2) arises from the following example. For simplicity of notation, the local coordinate systems  $(U_p, \varphi_p)$  and  $(U_q, \varphi_q)$  (chosen so that the above conditions are satisfied) are fixed and  $p' \in U_p$  is identified with  $\varphi_p(p') = (x_1, x_2) \in \varphi_p(U_p)$  and  $q' \in U_q$  is identified with  $\varphi_q(q') = (y_1, y_2) \in \varphi_q(U_q)$ .

Suppose the transformation  $\tau$  behaves as a dilation for all points  $p'$  in some small disc-like neighbourhood  $N_p$  of  $p$  ( $N_p \subset U_p$ ,  $\tau(N_p) \subset U_q$ ), so that

$$(x_1, x_2) \rightarrow (kx_1, kx_2), \quad (x_1, x_2) \in N_p, k > 1.$$

Given an object (or part of an object)  $A \in F(N_p)$  ( $N(A)$  not empty), let the associated luminous flux be considered a conserved quantity. Under the action of  $\tau$ , this fixed amount of flux must then be spread over the enlarged area of  $N(\tau(A))$ . It follows that the illumination  $(\tau(A))(q')$  at each point  $q' \in N(\tau(A))$  is therefore reduced by the factor  $1/k^2 = 1/\det(D\tau(p))$ .

More generally, for any  $\tau \in \text{Inj}(U, V)$ , the area of the image  $\tau(N_p)$  of the disc-like neighbourhood  $N_p$  is given by

$$\int_{\tau(N_p)} dy_1 dy_2 = \int_{N_p} |\det(D\tau(x_1, x_2))| dx_1 dx_2$$

The "dilation ratio" at  $p$  is then obtained by taking the limit:

$$\lim_{\substack{\text{diam}(N_p) \rightarrow 0 \\ p \in N_p}} \left| \frac{\int_{\tau(N_p)} dy_1 dy_2}{\int_{N_p} dx_1 dx_2} \right| = |\det(D\tau(p))|$$

Which, if either, of the definitions (1) or (2) is more appropriate depends upon the physical method by which the transformation  $\tau$  is effected. For the sake of brevity, (1) will be applied throughout the subsequent discussion.

Using the notation set out above, we now give a more precisely formulated version of the problem discussed in the introduction.

**Problem 1.** Given an object  $A$  and some transform  $\tau(A)$  of  $A$ , for what transformations  $\tau$  can the visual system recognise  $\tau(A)$  to be the original object  $A$ ?

(Such transformations  $\tau$  will be referred to as *invariance transformations*.)

Since our interest is oriented towards the construction of neural mechanisms for invariant recognition, we put Problem 1 in a more general form better suited to these ends. We will regard the visual recognition of an object  $B$  (as some object  $A$ ) as an effectuation, by the visual system, of a certain one-to-one correspondence (transformation) between  $B$  and  $A$  (and will refer to the visual system *executing, effecting*

or performing this transformation, and, in general, of its carrying this transformation).

**Problem 2.** Given an object  $A$ , what is the structure of the set of transformations which the visual system can perform upon  $A$ ?

In the ensuing, a method is developed by which, it is proposed, Problem 2 (and therefore Problem 1, also) may be answered – at least in part. The approach is founded upon a visual apparent motion effect the *phi-phenomenon* (Wertheimer, 1912):

Suppose the visual system is presented with a spot of light,  $S$  say, first in retinal location  $a$ , and second in retinal location  $b$ . Then, depending on parameters like the size and intensity of  $S$ , the locations  $a$  and  $b$ , the presentation times at  $a$  and  $b$ , and the time lag between the presentations, a visual sensation of motion can be induced in which the spot  $S$  appears to move across from  $a$  to  $b$ .

If we represent  $S$  at location  $a \in V$  by the mapping  $S_a: V \rightarrow R$ , and at location  $b \in V$  by the mapping  $S_b: V \rightarrow R$ , then the appearance of motion between locations  $a$  and  $b$  is nothing other than the ordered production of the objects  $\sigma_t(S_a)$ ,  $t \in [0, t_1]$ , where the  $\sigma_t$  are translations defined locally on the retina such that  $\sigma_0(S_a) = S_a$  and  $\sigma_{t_1}(S_a) = S_b$ . The time parametrization  $t \rightarrow \sigma_t$  is required to be reasonably smooth (differentiable).

Phi-motion can be generated between other than objects and their translates. [Rotations of an object are explicitly considered in the paper (Foster, 1972 b) which follows this.] We thus want a more generalized form of the above statement with  $S_a$  replaced by  $A \in F(U)$  and  $\sigma_t$  replaced by  $\tau_t \in \text{Inj}(U, V)$ . We will need the notions of a 1-parameter family of local transformations, a local 1-parameter monoid (and group) of local transformations, and the orbit of an object under each of these systems. The first three are either standard or obvious items, and their definitions are given in the Appendix. The last, the orbit of an object, is defined below. Note that we regard the orbit as a mapping into the set of all objects defined locally on  $V$ , rather than as a subset of this set. Note also that we adopt the convention of writing  $\psi_t$ ,  $t \in [t_1, t_2]$ , or just  $\psi_t$ , for the 1-parameter family of local transformations  $\psi: [t_1, t_2] \times U \rightarrow V$ .

The orbit  $\Omega(A)$  of an object  $A \in F(U)$  under the action of a 1-parameter family of local transformations  $\psi_t: U \rightarrow V$ ,  $t \in [t_1, t_2]$  is defined as the mapping

$$\Omega(A): t \in [t_1, t_2] \rightarrow \psi_t(A) \in \bigcup \{F(U'): U' \subset V\}$$

where it is recalled that the action of a transformation  $\psi_t \in \text{Inj}(U, V)$  on an object  $A \in F(U)$  is defined by setting  $(\psi_t(A))(p) = A(\psi_t^{-1}(p))$ .

With the aid of the above, we now proceed with the generalized formulation of the phi-phenomenon asked for earlier.

**Phi-Phenomenon (generalized).** Suppose the visual system is presented first with some object  $A \in F(U)$  and second with some transform  $\tau(A)$  of  $A$  ( $\tau \in \text{Inj}(U, V)$ ), then, depending upon  $A$ ,  $\tau(A)$ , exposure times, and time-lag between exposures, the orbit of  $A$  under a certain 1-parameter family of local transformations  $\psi_t: U \rightarrow V$ ,  $t \in [0, t_1]$ , is generated. The mapping  $\psi$  satisfies the conditions:

$$\psi_0(p) = p \quad \text{and} \quad \psi_{t_1}(p) = \tau(p) \quad \text{for all } p \in A.$$

It is emphasized that the family need not be a local monoid.

Obviously, a necessary condition for the existence of phi-motion is that the system be capable of performing transformations such as  $\psi_t$ , and, in particular, be capable of performing the transformation  $\psi_{t_1}$ . We can thus assert a partial solution to Problem 1.

**Proposition 1.** A transformation  $\tau \in \text{Inj}(U, V)$  is an invariance transformation (in the sense of Problem 1) for an object  $A \in F(U)$  if phi-motion can be induced between  $A$  and  $\tau(A)$ .

Note that we do not (yet) claim that the converse of this holds.

Turning to Problem 2, we immediately have, by the same reasoning that gives rise to Proposition 1, the following

**Proposition 2.** Given an object  $A \in F(U)$  and some transform  $\tau(A)$  ( $\tau \in \text{Inj}(U, V)$ ), the production of phi-motion between  $A$  and  $\tau(A)$ , so that for some family of local transformations  $\psi_t$ ,  $t \in [0, t_1]$ , the orbit  $\Omega(A)$  is generated, implies that the visual system is capable of performing a 1-parameter family of local transformations  $\psi'_t: U \rightarrow V$ ,  $t \in [0, t_1]$ , such that

$$\psi'_t(p) = \psi_t(p) \quad \text{for all } p \in A \quad \text{and} \quad t \in [0, t_1]. \quad (3)$$

For an appropriate choice of  $A$  and a sufficiently small neighbourhood  $U$  of  $N(A)$ , Eq. (3) serves to define  $\psi_t$ ,  $t \in [0, t_1]$ , to within visual indistinguishability and certain reparametrizations  $t \rightarrow \psi_t$ .

*Note.* As before, all that should strictly be asserted is that the visual system is capable of performing one out of a collection of mappings visually indistinguishable from  $\psi$  on  $A$ . In a full treatment, one would have tolerances on the sets  $U$ ,  $V$ ,  $[0, t_1]$ , and  $\text{Inj}(U, V)$ , and this ill-definedness would be naturally incorporated. To save further repetition of statements like the former, all subsequent assertions about sets, mappings,

etc., which relate to a visual process will be taken to be defined (in every case) not better than to within visual indistinguishability.

Suppose, now, that, for a given object  $A \in F(U)$  and some arbitrarily chosen transformation  $\tau \in \text{Inj}(U, V)$ , phi-motion is observed. By applying Proposition 2, we deduce that the visual system is capable of performing a certain 1-parameter family of local transformations  $\psi_t$ . With Problem 2 in mind, a natural step would be to consider the given family  $\psi_t$  in relation to the full set  $\text{Inj}(U, V)$ , the latter now with an appropriate differentiable structure defined upon it. In fact, the assigning of a differentiable structure to  $\text{Inj}(U, V)$  presents certain difficulties, and a more restricted approach must be adopted. So, instead of working with the full set  $\text{Inj}(U, V)$ , we limit our interest to certain subsets of  $\text{Inj}(U, V)$ , namely, those which carry the structure, locally, of a Lie transformation group. The definition of a local Lie group of local transformations.  $G \times U \rightarrow V$  is given in the Appendix. We make the additional stipulations that the local Lie group  $G$  is connected (equivalent to being path-connected), and the induced mapping  $\iota: g \in G \rightarrow \tau_g \in \text{Inj}(U, V)$  is injective. It follows from the last condition that any two distinct (differentiable) curves in  $G$

$$\gamma_i: t \in [t_1, t_2] \rightarrow \gamma_i(t) \in G, \quad i = 1, 2,$$

(which will be assumed to be injections) map into two distinct 1-parameter families of local transformations defined on  $U$ :

$$\tau_{\gamma_i(t)}, t \in [t_1, t_2], \quad i = 1, 2.$$

Similarly, any two distinct local 1-parameter subgroups in  $G$  map into two distinct local 1-parameter groups of local transformations defined on  $U$ . [For the definitions of a local 1-parameter subgroup (and submonoid), again see Appendix.]

Let us assume that, for a given object  $A \in F(U)$ , a suitable local Lie group of local transformations  $G \times U \rightarrow V$  has now been decided upon. We concentrate our attention on the local Lie group  $G$ , and ask the more restricted question: What subset  $G'$  of  $G \times U \rightarrow V$  can the visual system effect upon  $A$ ? (Compare with Problem 2.) To answer this, our procedure is clear. We select an element  $g$ , say, in  $G$ , and present the given object  $A$  and transform  $\tau_g(A)$  ( $\tau_g = \iota(g)$ ) in succession to the visual system. If phi-motion is induced between  $A$  and  $\tau_g(A)$ , then, by Proposition 2, we know the visual system can execute a certain 1-parameter family  $\psi_t$ ,  $t \in [0, t_1]$ . If  $\psi_t$  lies in  $\iota(G)$  for all  $t$  in  $[0, t_1]$ , then, by the preceding, there exists in  $G$  a path  $\gamma: [0, t_1] \rightarrow G$ , unique to within

extension, such that  $(\iota \circ \gamma)(t) = \psi_t$ , for all  $t \in [0, t_1]$ . Hence,  $G'$  contains, at least, the identity  $e$  (necessarily),  $g$ , and a path  $\gamma$  joining the two. We next choose a different element in  $G$ , say  $g'$  ( $g' \notin \gamma([0, t_1])$ ), and repeat the procedure. Continuing in this way, we build up an idea of the composition of (the identity component of)  $G'$ .

An important piece of information not considered in the above discussion is the possibility that the selected path  $\gamma$  is, in each case, part of a (local) 1-parameter subgroup in  $G$ . Apart from the general significance that 1-parameter subgroups have in relation to the structure of a Lie group, the selection, by the visual system, of such a path in preference to any other would have particularly important consequences in connection with the problem of constructing mechanisms for invariant recognition. Because of the subjective nature of the time-scale involved, however, it is difficult to obtain, by straightforward inspection, data upon the  $t$ -dependence of a given family  $\psi_t$ . Nevertheless, there is a less direct method, which, although failing to give the precise  $t$ -parametrization, does enable one to say whether  $\psi_t$  is a local 1-parameter monoid of local transformations. It depends upon the fact that for a given object  $A$ , we observe the orbit under  $\psi_t$  not of a single point  $p \in U$ , but of a large number of points, i.e., those in all of  $A$ .

Suppose we are given a 1-parameter family of local transformations  $\psi_t$ ,  $t \in [0, t_1]$ , acting upon an object  $A$ . We recall that  $\psi_t$  constitutes a local 1-parameter monoid of local transformations defined on  $A^1$  if for  $t, s, t+s \in [0, t_1]$ , and  $p, \psi_s(p) \in A$ , we have  $\psi_{t+s}(p) = \psi_t(\psi_s(p))$ . But this is the same as saying that, given any two points  $p$  and  $q$  in  $A$  such that  $q$  lies on the image of the orbit of  $p$ , after  $q$  the orbits of  $q$  and  $p$  coincide. By visual inspection of the way the object  $A$  is drawn along the orbit images of its elements, under the action of the family  $\psi_t$ , it should be possible to determine whether the required group property is exhibited. By way of illustration, let  $q_t, q'_t$ ,  $t \in [0, t_1]$ , be two 1-parameter families of rigid motions of the object  $A$  of Fig. 1 ( $V$  is identified with a hemisphere in  $\mathbf{R}^3$ ). For each  $t$  in  $[0, t_1]$ ,  $q_t$  and  $q'_t$  are determined by the images of any two distinct points in  $A$ . Consider the following two situations.

i) Let the family  $q_t$  act upon the object  $A$  in the manner indicated in Fig. 1 (i), so that the points  $p$  and  $q$  move through orbits  $a$  and  $b$  respectively. Let  $s \in [0, t_1]$  be such that  $q_s(p) = q$ . Define  $p_2 = q_{2s}(p)$ , and  $q_1 = q_s(q)$ . Then, clearly,  $q_{s+s}(p) \neq q_s(q_s(p))$ .

<sup>1</sup> Assume  $N(A)$  visually indistinguishable from some open  $U \supset N(A)$ .

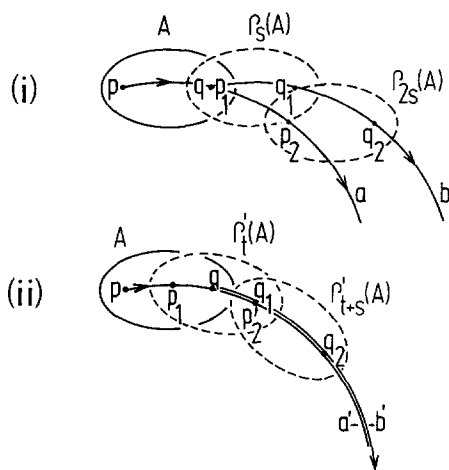


Fig. 1. The actions of the 1-parameter families of rigid motions  $q_t$  and  $q'_t$  upon the object A

ii) Let the family  $q'_t$  act upon A as in Fig. 1 (ii), so that the point p moves through the orbit a', and q moves through the orbit b', a' and b' coinciding after q. Let  $s \in [0, t_1]$  be such that  $q'_s(p) = q$ . Define  $q_1 = q'_t(q)$ , and  $p_2 = q'_{t+s}(p)$ . But by definition of a', b', we have  $p_2 = q_1$ . Hence,  $q'_{t+s}(p) = q'_t(q'_s(p))$ .

The conclusion follows that  $q_t$  is not a local 1-parameter monoid of local transformations, whereas  $q'_t$  is.

We thus have an experimental technique for indicating whether an observed family  $\psi_t$  corresponds (under  $\iota$ ) to a path  $\gamma$  which is a local 1-parameter submonoid in  $G$ . This method is now incorporated into the general investigatory procedure outlined earlier. The resulting program is set out in the next section.

### Investigatory Program

Before going ahead with the formulation of the proposed investigatory program, we recapitulate its aims and methods.

The intention is to discover something about the structure of the set of transformations which the visual system is capable of performing upon an object by virtue of the invariant recognition of the object under these transformations. A piecemeal approach is adopted in that rather than consider the whole set  $\text{Inj}(U, V)$  of all possible transformations of a subset  $U$  of the retina  $V$ , we limit ourselves at any one time to a selected local Lie group of local transformations  $G \times U \rightarrow V$ . For such a group, our interest then centres around determining how much of  $G$  is capable of being effected by the visual system. (This subset of  $G$  we denote by  $G'$ .)

The basis of the experimental technique is to take the generation of phi-motion between the given object and some selected transform as evidence that the system can execute a certain 1-parameter family of local transformations. By inspection of the orbit, one attempts to deduce whether the observed family is a local 1-parameter monoid of local transformations. This data is then related to the structure of  $G'$ .

We now proceed with the presentation of the program. Assume that the visual object  $A \in F(U)$  and local Lie group of local transformations  $G \times U \rightarrow V$  have already been selected.

(1) Choose an element  $g$  in  $G$ .

(2) Present the object A and transform  $\tau_g(A)$  in succession to the visual system.

(3) If phi-motion is induced, then check, by the method described in the theory, to see if the corresponding 1-parameter family of local transformations  $\psi_t, t \in [0, t_1]$ , is a local 1-parameter monoid of local transformations which can be factored through  $G$ :

$$\begin{array}{ccc} [0, t_1] & \xrightarrow{\psi_t} & \text{Inj}(U, V) \\ & \searrow \gamma & \nearrow \iota \\ & G & \end{array}$$

$\gamma: [0, t_1] \rightarrow G$  is a local 1-parameter submonoid in  $G \# (\gamma(0) = e, \gamma(t_1) = g)$ .

(4) Suppose such a  $\gamma$  exists [it is then determined uniquely up to extension by the diagram in (3)]. Form the local 1-parameter subgroup  $\gamma': I \rightarrow G \# ([0, t_1] \subset I \subset \mathbb{R}, I \text{ maximum})$  which coincides with  $\gamma$  on  $[0, t_1]$ .

(5) Select an element  $t$  from  $I$ , and present the object A and transform  $\tau_{\gamma'(t)}(A)$  in succession to the visual system.

(6) If phi-motion is induced, then check, as before, to see whether the corresponding family of transformations belongs to  $\gamma'$ .

(7) Repeat steps (5) and (6) until it is judged that a sufficiently dense collection of points  $t$  in  $I$  has been arrived at (dense in the sense that more points  $t$  would be unlikely to give rise to qualitatively new information about the system).

(8) Choose a different element  $g$  from  $G$ , where  $g \notin \gamma'(I)$ , and recommence the procedure from (2), unless, as in (7), a sufficiently dense collection of paths in  $G'$  has already been achieved.

For the chosen object and local Lie group of local transformations, this program is claimed both to yield data upon the composition of the identity component of  $G'$ , and to indicate when a path (in  $G'$ ) executed by the visual system coincides locally with a local 1-parameter subgroup in  $G'$ . In the paper immediately after this (Foster, 1972 b) we describe an experiment which shows the program in operation. A Landolt ring is

the visual object, and the group of rotations  $SO(2)$ , acting on the retina and about the centre of the ring, is the selected Lie group of local transformations.

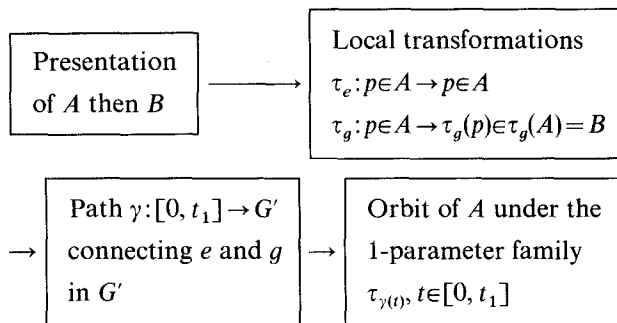
### Discussion

The method of analysis put forward in the two preceding sections does not necessarily give complete information on the structure of the set of transformations carried by the visual system, since, apart from the restriction of the method to connected local Lie groups of local transformations, there is the fact that Propositions 1 and 2 (in the theory) only go in one direction, i.e.,  $\phi$ -motion  $\Rightarrow$  invariance transformations.

In order to say something about the converses of Propositions 1 and 2, it is necessary to speculate upon possible systems by which  $\phi$ -motion (in the general form defined in the theory) can be generated. If we are prepared to limit the discussion to connected local Lie groups  $G$  only, then one fairly natural candidate is indeed a system which implies the converses to the two propositions.

Consider the following:

Built into the system, there is the capacity to operate upon visual objects with certain local transformations. Specifically, to each object  $A$  on the retina  $V$  there is a local Lie group of local transformations  $G' \times U \rightarrow V$  acting upon a (small) neighbourhood  $U$  of  $A$ .  $G'$  is not necessarily assumed connected. Under appropriate experimental conditions, successive presentation to the system of objects  $A$  and  $B$ , where  $B$  is such that there exists  $g \in G'$  for which  $\tau_g(A) = B$ , causes the system to attempt to select a path  $\gamma: [0, t_1] \rightarrow G'$  from  $e$  to  $g$  in  $G'$ . If  $\gamma$  exists, the system then operates upon the object  $A$  with the induced 1-parameter family of local transformations  $\tau_{\gamma(t)}$ ,  $t \in [0, t_1]$ , producing the transforms  $\tau_{\gamma(t)}(A)$  in order as  $t$  runs from 0 to  $t_1$ . Schematically,



If any of the mappings (arrows) is not defined, then  $\phi$ -motion is not obtained.

In the above, the phrase "appropriate experimental conditions" is used. This is meant to include things

like optimum exposure times, time-lag, intensity levels, etc. Given these, the essential requirement for the converse of Proposition 2 (and 1) to hold is that  $G'$  be path-connected; for the general case, this can only be established by independent experimentation.

*Acknowledgment.* The author is grateful to Dr. C. J. Isham for many useful discussions.

### Appendix

Let  $U$  be a (non-empty) open subset of the retina  $V$ ,  $[t_1, t_2]$  a closed interval in  $\mathbf{R}$ , and  $G$  a local Lie group [see, for example, Cohn (1968)].

(1) A 1-parameter family of local transformations defined on  $U$  is a differentiable mapping  $\psi$  of  $[t_1, t_2] \times U$  into  $V$  such that for each  $t \in [t_1, t_2]$ ,  $\psi_t: p \rightarrow \psi(t, p)$  is a diffeomorphism of  $U$  onto the open set  $\psi_t(U)$  of  $V$ .

(2) A local 1-parameter monoid (resp. group) of local transformations defined on  $U$  is a differentiable mapping  $\psi$  of  $[0, t_1] \times U$  (resp.  $[-t_1, t_1] \times U$ ) into  $V$  which satisfies:

i) For each  $t \in [0, t_1]$  (resp.  $[-t_1, t_1]$ ),  $\psi_t: p \rightarrow \psi(t, p)$  is a diffeomorphism of  $U$  onto  $\psi_t(U)$ .

ii) If  $t, s, t+s \in [0, t_1]$  (resp.  $[-t_1, t_1]$ ) and if  $p, \psi_s(p) \in U$ , then

$$\psi_{t+s}(p) = \psi_t(\psi_s(p)).$$

iii) For all  $p \in U$ ,  $\psi_0(p) = p$ .

(3) A local Lie group of local transformations defined on  $U$  is a differentiable mapping  $\tau: G \times U \rightarrow V$  which satisfies:

i) For each  $g \in G$ ,  $\tau_g: p \rightarrow \tau(g, p)$  is a diffeomorphism of  $U$  onto  $\tau_g(U)$ .

ii) If  $g_1, g_2, g_1 g_2 \in G$  and if  $p, \tau_{g_2}(p) \in U$ , then

$$\tau_{g_1 g_2}(p) = \tau_{g_1}(\tau_{g_2}(p)).$$

( $G$  acts on the left.)

iii) For all  $p \in U$ ,  $\tau_e(p) = p$  ( $e$  neutral in  $G$ ).

(4) A local 1-parameter submonoid (resp. subgroup) in  $G$  is a differentiable mapping  $\gamma$  of  $[0, t_1]$  (resp.  $[-t_1, t_1]$ ) into  $G$  such that if  $t, s, t+s \in [0, t_1]$  (resp.  $[-t_1, t_1]$ ), then the product  $\gamma(t)\gamma(s)$  is defined and

$$\gamma(t+s) = \gamma(t)\gamma(s).$$

### References

- Arbib, M.A.: How we know universals: retrospect and prospect. *Math. Biosci.* **11**, 95–107 (1971).  
 Cohn, P.M.: Lie groups. London: Cambridge Univ. Press 1968.  
 Foster, D.H.: A method for the investigation of those transformations under which the visual recognition of a given object is invariant: II. An example experiment: the group of rotations  $SO(2)$  acting on a Landolt ring. *Kybernetik* **11**, 223–229 (1972).  
 Pitts, W.H., McCulloch, W.S.: How we know universals – the perception of auditory and visual forms. *Bull. Math. Biophys.* **9**, 124–147 (1947).  
 Wertheimer, M.: Experimentelle Studien über das Sehen von Bewegung. *Z. Psychol.* **61**, 161–265 (1912).  
 Zeeman, E.C.: The topology of the brain and visual perception. In: The topology of 3-manifolds and related topics. New Jersey: Prentice-Hall 1962.

Dr. D. H. Foster  
 Imperial College of Science and  
 Technology  
 Dept. of Physics  
 Prince Consort Road  
 London SW7 2BZ, Great Britain